SQL Capstone Project - Analytical CRM Development for a Bank

Note: 1.all the tables used for reference here were derived from SQL queries and queries can be found in SQL text file follow by commented questions with same question number mentioned here for additional info.

2.all charts and visuals mentioned here can be found in power bi file for viewing and analysis purposes.

Objective Questions

Q1.What is the distribution of account balance across different regions?

Ans: The account balance were ₹31.13 CRs, ₹30.04 CRs, ₹15.31 CRs for France, Germany, Spain respectively as per following table

|  |  |
| --- | --- |
| regions | Account balance |
| France | ₹ 31,13,32,479.49 |
| Germany | ₹ 30,04,02,861.38 |
| Spain | ₹ 15,31,23,552.01 |

Refer chart name distribution of account balance across different regions for an visual

Q2.Identify the top 5 customers with the highest number of transactions in the last quarter of the year. (SQL)

Ans: top 5 customers with the highest number of transactions in the last quarter of the latest year in database were as following

|  |  |  |
| --- | --- | --- |
| Customer Id | Customer name | transactions |
| 15608528 | Munro | 4 |
| 15625824 | Kornilova | 4 |
| 15633194 | Osborne | 4 |
| 15634974 | Seppelt | 4 |
| 15668889 | Galgano | 4 |

Note: as there more than 5 customers with 4 transactions present in the last quarter of the latest year in database hence results were sorted with transactions in descending, customer id in ascending, customer name alphabetically and with limited it 5 to show only 5 customers

Q3.Calculate the average number of products used by customers who have a credit card. (SQL)

Ans: the average number of products used by customers who have a credit card is 2.

Q4.Determine the churn rate by gender for the most recent year in the dataset.

Ans: the churn rate were 25.05%,15.37% for female, male respectively for most recent year in the dataset as per following table

|  |  |
| --- | --- |
| Gender Category | Churn rate |
| Female | 25.05% |
| Male | 15.37% |

Refer chart name churn rate by gender for an visual.

Q5.Compare the average credit score of customers who have exited and those who remain. (SQL)

Ans: the average credit score were 645.35, 651.85 for exited customers, retained customers respectively as per following table

|  |  |
| --- | --- |
| Customer Category | Avg credit score |
| Exited customers | 645.35 |
| Retained customers | 651.85 |

Q6.Which gender has a higher average estimated salary, and how does it relate to the number of active accounts? (SQL)

Ans: by referring below table we can say that there is inverse co-relation between average estimated salary and number of active accounts as per data we can see that female customers have higher average estimated salary then male customers but female customers have lower active accounts then male customers .

|  |  |  |
| --- | --- | --- |
| Gender Category | Avg EstimatedSalary | Active accounts |
| Female | ₹ 1,00,601.54 | 2284 |
| Male | ₹ 99,664.58 | 2867 |

Q7.Segment the customers based on their credit score and identify the segment with the highest exit rate. (SQL)

Ans: by referring below table we can say that the segment with the highest exit rate is 850 credit score Segment with 10.24% exit rate.

|  |  |
| --- | --- |
| Credit Score | Exit rate |
| 850 | 10.24% |
| 651 | 4.05% |
| 705 | 3.81% |
| 637 | 3.33% |
| 678 | 3.1% |

Q8.Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)

Ans: by referring below table we can say that geographic region that has highest number of active customers with a tenure greater than 5 years is France with number of active customers of 797.

|  |  |
| --- | --- |
| regions | Active customers |
| France | 797 |
| Spain | 431 |
| Germany | 399 |

Q9.What is the impact of having a credit card on customer churn, based on the available data?

Ans: by referring below table we can say that the impact of having a credit card is having negative impact towards retaining customers.

|  |  |
| --- | --- |
| Category | Exited customers |
| credit card holder | 1424 |
| non credit card holder | 613 |

Refer chart name Customer churn by credit card status for an visual

Q10.For customers who have exited, what is the most common number of products they had used?

Ans: by referring below table we can say that the most common number of products Exited customers had used is 1.

|  |  |
| --- | --- |
| Num Of Products | Exited customers |
| 1 | 1409 |
| 3 | 220 |
| 2 | 348 |
| 4 | 60 |

Refer chart name number of products used by exited customers for an visual

Q11.Examine the trend of customer exits over time and identify any seasonal patterns (yearly or monthly). Prepare the data through SQL and then visualize it.

Ans: by referring below table we can say that the trend of customer exits over time is increasing year by year.

|  |  |  |  |
| --- | --- | --- | --- |
| year | churn | Pre churn | trend |
| 2016 | 376 | 0 | increasing |
| 2017 | 479 | 376 | increasing |
| 2018 | 524 | 479 | increasing |
| 2019 | 658 | 524 | increasing |

Refer chart name the trend of customer exits over time for an visual

Q12.Analyse the relationship between the number of products and the account balance for customers who have exited.

Ans: after analysing scatter chart we can say that there is no relationship between the number of products and the account balance for Exited customers.

Refer chart name Relation between no. of products and the account balance of exited customers for an visual

Q13.Identify any potential outliers in terms of spend among customers who have remained with the bank.

Ans: No, we can’t Identify any potential outliers in terms of spend among retained customers, because there is no data point that differs significantly from other observations.

Q14.Can you create a dashboard incorporating the visuals mentioned above and additionally derive more KPIs if possible?

Ans: Yes, we can create a dashboard incorporating the visuals mentioned and we can include KPIs also

Refer report name main Objective Questions visuals for dashboard

Q15.Using SQL, write a query to find out the gender wise average income of male and female in each geography id. Also rank the gender according to the average value. (SQL)

Ans: the gender wise average income of male and female in each geography with along respective rank in each geography were as following

|  |  |  |  |
| --- | --- | --- | --- |
| **Geography Location** | **Gender Category** | **Average income** | **ranks** |
| France | Male | 100174.25 | 1 |
| France | Female | 99564.25 | 2 |
| Germany | Female | 102446.42 | 1 |
| Germany | Male | 99905.03 | 2 |
| Spain | Female | 100734.11 | 1 |
| Spain | Male | 98425.69 | 2 |

Q16.Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).

Ans: the average tenure of the people who have exited in each age bracket were as following

|  |  |
| --- | --- |
| **Age bracket** | **Average tenure** |
| 31-50 | 4.87 |
| 18-30 | 4.84 |
| 50+ | 4.85 |

Q**17**.Is there any direct correlation between salary and balance of the customers? And is it different for people who have exited or not?

Ans: after analysing scatter chart we can say that there is no direct correlation between salary and balance of the customers and it is same for both retained customers and exited customers.

Refer chart name relation between salary and balance of retained customers and relation between salary and balance of exited customers for an visual

Q**18**. Is there any correlation between salary and Credit score of customers?

Ans: after analysing scatter chart we can say that there is no correlation between salary and Credit score of customers.

Refer chart name relation between salary and Credit score of customers for an visual

Q19.Rank each bucket of credit score as per the number of customers who have churned the bank.

Ans: please find below table with bucket of credit score along with their number of exited customers and their ranks respectively.

|  |  |  |
| --- | --- | --- |
| **Credit score bucket** | **churn** | **ranks** |
| 580–669 | 685 | 1 |
| 300–579 | 520 | 2 |
| 670–739 | 452 | 3 |
| 740–799 | 252 | 4 |
| 800–850 | 128 | 5 |

Q20.According to the age buckets find the number of customers who have a credit card. Also retrieve those buckets who have lesser than average number of credit cards per bucket.

Ans1: please find below table with age buckets along with their credit card holders.

|  |  |
| --- | --- |
| **Age bracket** | **Credit card holders** |
| 18-30 | 1400 |
| 31-50 | 4781 |
| 50+ | 874 |

Ans2: please find below table with age buckets along with their credit card holders who have lesser than average number of credit cards per bucket

|  |  |
| --- | --- |
| **Age bracket** | **Credit card holders** |
| 18-30 | 1400 |
| 50+ | 874 |

Q21.Rank the Locations as per the number of people who have churned the bank and average balance of the learners.

Ans: please find below table with Locations along with their number of exited customers, average balance and ranks respectively.

|  |  |  |  |
| --- | --- | --- | --- |
| **Geography Location** | **Exited customers** | **Avg balance** | **ranks** |
| Germany | 814 | 120361.08 | 1 |
| France | 810 | 71192.8 | 2 |
| Spain | 413 | 72513.35 | 3 |

Subjective Question

1.Customer Behaviour Analysis:

Q1.What patterns can be observed in the spending habits of long-term customers compared to new customers, and what might these patterns suggest about customer loyalty?

Ans: by referring below table we can be observe there is no significance difference in Spending habits of long-term customers compared to new customers and hence it is difficult to suggest about customer loyalty based on these patterns.

|  |  |  |
| --- | --- | --- |
| **Tenure category** | **Num Of Products per customers** | **Customers percentage** |
| long-term customers | 1.51 | 31.71% |
| medium-term customers | 1.53 | 54.81% |
| new customers | 1.56 | 13.48% |

note: long-term customers are those customers who Tenure is greater than 5 years,

medium-term customers are those customers who Tenure is between 4 and 5 years,

new customers are those customers who Tenure is lesser than or equal to 3 years,

those definition were created after finding minimum Tenure of customers (3 years), average Tenure of customers (5 years) and maximum Tenure of customers (7 years)

2.Product Affinity Study:

Q2.Which bank products or services are most commonly used together, and how might this influence cross-selling strategies?

Ans: there is lack of data in dataset to find bank products or services are most commonly used together as there is only one data point exist (Num Of Products) the indicates customers spending, purchase, products or services usage and transactions made by them and hence we can’t make any cross-selling strategies

3.Geographic Market Trends:

Q3.How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?

Ans: by referring below table we can say that there is no correlation with the number of active accounts and customer churn rates in different geographic regions.

|  |  |  |
| --- | --- | --- |
| **Geography Location** | **Active customers percentage** | **Exited customers percentage** |
| France | 50.3 | 39.76 |
| Spain | 25.47 | 20.27 |
| Germany | 24.23 | 39.96 |

Refer chart name Active customers and exited customers by Geography Location for an visual

4.Risk Management Assessment:

Q4.Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?

Ans: by referring below table we can say that there is no particular demographic segments appears to pose the highest financial risk to the bank based on average Credit Score

|  |  |  |  |
| --- | --- | --- | --- |
| **Geography Location** | **Gender Category** | **Age bracket** | **Avg CreditScore** |
| Germany | Female | 18-30 | 650.56 |
| France | Male | 18-30 | 654.78 |
| Spain | Male | 18-30 | 649.76 |
| France | Female | 18-30 | 650.24 |
| Spain | Female | 18-30 | 644.48 |
| Germany | Male | 18-30 | 652.48 |
| France | Female | 31-50 | 649.27 |
| Spain | Female | 31-50 | 652.58 |
| Spain | Male | 31-50 | 651.14 |
| France | Male | 31-50 | 648.19 |
| Germany | Male | 31-50 | 649.86 |
| Germany | Female | 31-50 | 656.2 |
| Germany | Male | 50+ | 646.95 |
| France | Female | 50+ | 646.99 |
| Spain | Female | 50+ | 658.41 |
| France | Male | 50+ | 652.48 |
| Spain | Male | 50+ | 652.11 |
| Germany | Female | 50+ | 642.28 |

5.Customer Lifetime Value Forecast:

Q5.How would you use the available data to model and predict the lifetime value of different customer segments?

Ans: by referring below table we can say that customers of age range between 31 to 50 can be consider as most valuable customer segment their total account balance Significantly higher than other customer segments and can used for most banking activities for business growth.

|  |  |  |
| --- | --- | --- |
| **Age bracket** | **Account balance** | **Avg CreditScore** |
| 31-50 | ₹51,99,85,034.96 | 650.5 |
| 18-30 | ₹14,40,55,167.65 | 651.2 |
| 50+ | ₹10,08,18,690.27 | 649.63 |

6.Marketing Campaign Effectiveness:

Q6.How could you assess the impact of marketing campaigns on customer retention and acquisition within the dataset?

Ans: there is no data point related to marketing campaigns made by bank, hence because of lack of data we can’t assess the impact of marketing campaigns on customer retention and acquisition.

7.Customer Exit Reasons Exploration:

Q7.Can you identify common characteristics or trends among customers who have exited that could explain their reasons for leaving?

Ans: After analysing all possible Customer Exit Reasons we can find top 3 common characteristics among customers who have exited and this top 3 common characteristics is based on positive correlation between exited customers and churn rates ,those 3 top common characteristics were as following

1.Inactive members/customers : with 26.85% churn rate among Inactive members which resulted 1302 number of customers to exit.

2.Female customers : with 25.07% churn rate among Female customers which resulted 1139 number of customers to exit.

3.Germany region : with 32.44% churn rate in Germany region which resulted 814 number of customers to exit.

Refer report name Main Subjective Questions visuals for visuals

Q8.Are 'Tenure', 'Num Of Products', 'Is Active Member', and 'Estimated Salary' important for predicting if a customer will leave the bank?

Ans: After analysing data we can say that Num Of Products, Is Active Member is important for predicting if a customer will leave the bank, but Tenure, Estimated Salary is not that important for predicting if a customer will leave the bank because those 2 data point churn rate is similar.

Refer report name Main Subjective Questions visuals for visuals

Q9.Utilize SQL queries to segment customers based on demographics, account details, and transaction behaviours.

Ans1: please find below table with customers segments based on customers demographics.

|  |  |  |  |
| --- | --- | --- | --- |
| **Geography Location** | **Gender Category** | **Age bracket** | **Customers count** |
| France | Male | 31-50 | 1861 |
| France | Female | 31-50 | 1517 |
| Spain | Male | 31-50 | 954 |
| Germany | Male | 31-50 | 897 |
| Germany | Female | 31-50 | 803 |
| Spain | Female | 31-50 | 739 |
| France | Male | 18-30 | 576 |
| France | Female | 18-30 | 462 |
| France | Male | 50+ | 316 |
| France | Female | 50+ | 282 |
| Spain | Male | 18-30 | 267 |
| Germany | Male | 18-30 | 245 |
| Spain | Female | 18-30 | 210 |
| Germany | Female | 18-30 | 208 |
| Germany | Female | 50+ | 182 |
| Germany | Male | 50+ | 174 |
| Spain | Male | 50+ | 167 |
| Spain | Female | 50+ | 140 |

Ans2: please find below table with customers segments based on customers account details.

|  |  |  |  |
| --- | --- | --- | --- |
| **Joining year** | **Tenure** | **Balance range** | **Customers count** |
| 2019 | 4 | 75001-150000 | 990 |
| 2019 | 4 | 0-75000 | 791 |
| 2018 | 5 | 75001-150000 | 759 |
| 2019 | 3 | 75001-150000 | 696 |
| 2018 | 5 | 0-75000 | 662 |
| 2017 | 6 | 75001-150000 | 604 |
| 2016 | 7 | 75001-150000 | 594 |
| 2019 | 3 | 0-75000 | 522 |
| 2018 | 4 | 75001-150000 | 501 |
| 2017 | 6 | 0-75000 | 474 |
| 2017 | 5 | 75001-150000 | 450 |
| 2016 | 7 | 0-75000 | 447 |
| 2018 | 4 | 0-75000 | 409 |
| 2016 | 6 | 75001-150000 | 396 |
| 2017 | 5 | 0-75000 | 390 |
| 2016 | 6 | 0-75000 | 346 |
| 2019 | 4 | 150001-225000 | 184 |
| 2018 | 5 | 150001-225000 | 166 |
| 2017 | 6 | 150001-225000 | 142 |
| 2019 | 3 | 150001-225000 | 130 |
| 2016 | 7 | 150001-225000 | 98 |
| 2018 | 4 | 150001-225000 | 95 |
| 2017 | 5 | 150001-225000 | 82 |
| 2016 | 6 | 150001-225000 | 70 |
| 2017 | 5 | 225000+ | 1 |
| 2018 | 4 | 225000+ | 1 |

Ans3: please find below table with customers segments based on customers transaction behaviours.

|  |  |
| --- | --- |
| **NumOf Products** | **Customers count** |
| 1 | 5084 |
| 2 | 4590 |
| 3 | 266 |
| 4 | 60 |

Q10.How can we create a conditional formatting setup to visually highlight customers at risk of churn and to evaluate the impact of credit card rewards on customer retention?

Ans1: we can create using conditional formatting feature and assigning various colours in conditional formatting for different values in power bi to highlight customers at risk of churn

Refer table name Conditional Formatted Table for an visual

Ans2: we can’t create conditional formatting to evaluate the impact of credit card rewards on customer retention because there no data point that related to credit card rewards in database.

Q11.What is the current churn rate per year and overall as well in the bank. Can you suggest some insights to the bank about which kind of customers are more likely to churn and what are the different strategies that can be used to decrease the churn rate.

Ans1: The current churn rate per year and overall as well in the bank were as following.

|  |  |  |
| --- | --- | --- |
| **year** | **yearly churn rate** | **overall churn rate** |
| 2016 | 19.27% | 20.37% |
| 2017 | 22.35% | 20.37% |
| 2018 | 20.21% | 20.37% |
| 2019 | 19.86% | 20.37% |

Ans2: After analysing data we can say that the customers who have aged 51+ years old, the customers with account balance greater than or equal to 225001 and The customers who are using 3 or more products have highest churn rates other kind of customers and those kind of customers are more likely to churn.

For the customers who have aged 51+ years old we can introduce senior citizens benefits plan or we can do similar activities to decrease the churn rate.

For the customers with account balance greater than or equal to 225001 we can suggest various investment plan with competitive rates for their funds to help them have future secured financial status for better living to decrease the churn rate.

For The customers who are using 3 or more products we can provide considerable rates of discounts by encouraging them use more products for better prices to decrease the churn rate.